

TIBBIYOT MASALALARINIYK KOMPLEKS TADQIQIDA INFORMATIV SIMPTOKOMPLEKSLARNI SHAKLLANTIRISH ALGORITMI

Axram Xasanovich NISHANOV

professor

texnika fanlari doktori

Toshkent axborot texnologiyalari universiteti

Toshkent, O'zbekiston

nishanov_akram@gmail.ru

G'ulomjon Primovich JO'RAYEV

texnika fanlari bo'yicha (PhD) falsafa doktori

Qashqadaryo viloyati pedagoglarni yangi

metodikalarga o'rgatish milliy markazi

Qarshi, O'zbekiston

g.p.jurayev@mail.ru

Saidqul Xo'jamurodovich SAPAROV

doktorant

Toshkent axborot texnologiyalari universiteti

Toshkent, O'zbekiston

saidqul80@bk.ru

Annotatsiya

Maqolada tibbiyot sohasidagi Faringit va Gepatit kasalliklari uchun shakllantirilgan o'quv tanlamalariga dastlabki ishlov berish masalalarining yechishga e'tibor qaratilgan. Bunda dastlab o'quv tanlama obyektlarining belgi qiymatlari normallashtirish va belgi qiymatlarini 0 yoki 1 qiymatga o'tkazish, sinf obyektlarini sinflashtirish orqali etalon o'quv tanlamani shakllantirish, informativ belgilar majmualarini tanlash, tanlangan informativ belgilar majmualarining sinflashtirish samaradorligini baholash hamda informativ belgilar majmualari kesimida informativ simptokomplekslarni shakllantirish kabi masalalarni yechish algoritmlari taklif etilgan.

Tayanch so'zlar: o'quv tanlama, dastlabki ishlov berish, normallashtirish, informativ belgilar majmuasi, informativ simptokomplekslar, sinflashtirish.

АЛГОРИТМ ФОРМИРОВАНИЯ ИНФОРМАТИВНЫХ СИМПТОКОМПЛЕКСОВ В КОМПЛЕКСНОМ ИССЛЕДОВАНИИ МЕДИЦИНСКИХ ВОПРОСОВ

Ахрам Хасанович НИШАНОВ

профессор

доктор технических наук

Ташкентский университет информационных технологий

Ташкент, Узбекистан

Гуломжон Примович ЖУРАЕВ

доктор философии (PhD) по техническим наукам

Национальный центр подготовки учителей

новым методам Кашкадарьинской области

Карши, Узбекистан
Саидкул Хужамуродович САПАРОВ

докторант
Ташкентский университет информационных технологий
Ташкент, Узбекистан

Аннотация

В статье рассмотрены вопросы первичной обработки обучающей выборки, сформулированной для больных фарингитом и гепатитом в сфере медицины. При этом изначально были предложены алгоритмы решения таких задач, как нормализация пространств признаков объектов обучающей выборки и преобразование пространств признаков в значения 0 или 1, формирование эталонной выборки путём классификации объектов класса, отбора информативных наборов признаков, оценивая эффективность классификации выбранных информативных наборов признаков и формируя информативные симптомокомплексы в контексте информативных наборов признаков.

Ключевые слова: обучающая выборка, первичная обработка, нормализация, информативных наборов признаков, информативные симптомокомплексы, классификация.

Jahonda barcha sohalarda bo'lgani singari sog'liqni saqlash sohasini tizimli rivojlantirishda axborot-kommunikatsiya texnologiyalarini jadal sur'atlarda joriy etish orqali aholiga tibbiy yordam ko'rsatish, kasallik turlarini va yuzaga kelish sabablarini erta aniqlash hamda maqsadli davolash usullarini takomillashtirish borasida olib boriladigan ishlarga alohida e'tibor qaratilmoqda. Ayniqsa, hozirgi vaqtga kelib tibbiy ma'lumotlar hajmining keskin ortib borgani sayin, ularni intellektual tahlil qilish tizimlarini ishlab chiqish va amaliyotga tadbiq etish muhim ahamiyat kasb etmoqda. Ayniqsa, tibbiy ma'lumotlarni intellektual tahlil qilish masalalarini hal etishda tadqiqot obyektlarini xarakterlovchi belgilar va belgilar majmuasi aro yashirin bog'liqliklarni hamda qonuniyatlarni aniqlashda o'lcham, fazo muammolariga duch kelinadi. Odatda, kasallik holatini baholovchi, yig'ilgan va saqlanayotgan tibbiy ma'lumotlar o'lchami o'nlab yoki yuzlab belgilardan iborat bo'ladi. Ushbu ma'lumotlarga asoslanib, soha mutaxassislari tomonidan qaror qabul qilish o'ta murakkab jarayonga aylanadi. Shu bois, tibbiy ma'lumotlarga ishlov berishda katta o'lchamdan kichik muhimroq o'lchamga o'tish, ya'ni informativ belgilar majmualarini tanlash va baholash hamda

informativ simptokomplekslarni shakllantirish va ular asosida tibbiy timsollarni tashxislash, shuningdek, soha mutaxassislariga ko‘mak beruvchi zamonaviy tanib olish tizimlarni ishlab chiqish hozirgi kunda sog‘liqni saqlash tizimining dolzarb masalalaridan biri hisoblanadi.

Shu bois, amaliy tibbiyot masalalari yechishda soha mutaxassislari tomonidan shakllantirilgan o‘quv tanlama sifatini baholash va ular asosida ma’lumotlarga dastlabki ishlov berishning usul, algoritmlarini ishlab chiqish, mavjudlarini takomillashtirish borasida ilmiy izlanishlar olib borilmoqda, masalan, o‘quv tanlanma ma’lumotlarini tahlil va integratsiya qilish [1-6-6], o‘quv tanlanma obyektlarini parametrlari kesimida normallashtirish [7-8-9], o‘quv tanlanma asosida etalon jadvalni shakllantirish [10-11-12], sinflashtirish, obyektlarni tasniflab beruvchi informativ belgilar majmuasini tanlash, informativ belgilar fazosida sinf obyektlari, belgilar va sinflarning muhimlik darajalarini aniqlash [13-14].

Maqolada tibbiyot sohasidagi model masalalar sifatida Faringit va Gepatit kasalliklari uchun shakllantirilgan o‘quv tanlamalar berilgan bo‘lib, ular ustida tajribaviy tadqiqot ishlari o‘tkazildi va natijalar olindi. Dastlab, Faringit va Gepatit kasalliklari uchun shakllantirilgan o‘quv tanlamalarning umumiy xarakteristikasi hamda o‘quv tanlamalardagi obyektlarini xarakterlovchi belgilarning nomlanishi quyidagi 1–2-jadvallarda ifoda etilgan.

1-jadval

Modelli masalalarning umumiy xarakteristikalari

Masalaning nomi	Sinflar soni	Belgilar soni	Obyektlar soni
Faringit	2	18	676
Gepatit	2	19	155

2-jadval

Belgilarning nomlanishi

Faringit	Gepatit
x^1 -yoshi(age);	x^1 -yoshi(age);
x^2 -og‘riq(pain);	x^2 -jinsi(sex);
x^3 -shishgan(swollenadp);	x^3 -steroid(steroid);
x^4 - tender(tender);	x^4 -virusga qarshi(antivirals);

x^5 -bodomsimon bezlar shishi (tonsillarswelling);	x^5 -charchoq(fatigue);
x^6 -ekssudat(exudate);	x^6 -bezovtalik(malaise),
x^7 -harorat(temperature);	x^7 -anoreksiya(anorexia);
x^8 -to'satdan(sudden);	x^8 -katta jigar(liver big);
x^9 -yo'tal(cough);	x^9 -qattiq jigar(liver firm);
x^{10} -burun oqishi(rhinorrhea);	x^{10} -taloq palpatsiyasi(spleen palpable);
x^{11} -konyunktivit(conjunctivitis);	x^{11} -o'rgimchaklar(spiders);
x^{12} -bosh og'rig'i(headache);	x^{12} -assit(ascites);
x^{13} -eritema(erythema);	x^{13} -varikoz(varices);
x^{14} -petexiya(petechiae);	x^{14} - bilirubin (bilirubin);
x^{15} -qorin og'rig'i(abdopain);	x^{15} -ishqoriy fosfat(alk phosphate);
x^{16} -diareya(diarrhea),	x^{16} -SGOT(сывороточная глутаминовая щавелевоуксусная трансаминаза) (SGOT);
x^{17} -qayt qilish(nauseavomit);	x^{17} -albumin(albumin);
x^{18} -qizarish(scarlet).	x^{18} -taxminiyl vaqt(protime);
	x^{19} -gistologiya(histology).

a) masalaning qo'yilishi. Faraz qilaylik, o'quv tanlanmalar sinflarga ajratilgan va ular quyidagicha berilgan bo'lsin:

$$X_1 = \begin{bmatrix} x_{11}^1 & x_{11}^2 & \dots & x_{11}^N \\ x_{12}^1 & x_{12}^2 & \dots & x_{12}^N \\ \vdots & \vdots & \vdots & \vdots \\ x_{1m_p}^1 & x_{1m_p}^2 & \dots & x_{1m_p}^N \end{bmatrix} \dots X_r = \begin{bmatrix} x_{r1}^1 & x_{r1}^2 & \dots & x_{r1}^N \\ x_{r2}^1 & x_{r2}^2 & \dots & x_{r2}^N \\ \vdots & \vdots & \vdots & \vdots \\ x_{rm_r}^1 & x_{rm_r}^2 & \dots & x_{rm_r}^N \end{bmatrix}.$$

bu yerda $p = \overline{1, r}$; hamda o'quv tanlanma $X = \bigcup_{p=1}^r X_p$ ko'rinishda ifodalanib, ular o'zaro kesishmaydigan sinflardan iborat va quyidagi ya'ni $X_p \cap X_q = \emptyset, (p, q = \overline{1, r}; p \neq q.)$ shartlar berilgan bo'lsin.

Xuddi shuningdek, obyekt x_{pi} -ning komponentalari x_{pi}^j -haqiqiy sonlardan iborat. Bu yerda $p = \overline{1, r}; i = \overline{1, m_p}; j = \overline{1, N}$; hamda r -berilgan sinflarning umumiy soni, $m_p - p$ -sinfdagi bemorlarning umumiy soni va N -belgilarning umumiy sonini bildiradi.

b) quyidagi masalalarni yechish talab etiladi:

- 1) X_p sinf obyektlarini xarakterlovchi belgi qiymatlarni normallashtirish;
- 2) X_p sinflardagi obyektlarning belgi qiymatlari uzluksiz miqdordan 0 yoki 1 qiymatga o'tkazish;
- 3) X_p sinf obyektlarini o'z sinfini shakllanishiga qo'shgan hissasini baholash orqali etalon o'quv tanlamasini shakllantirish;

- 4) informativ belgilar majmualarini tanlash;
- 5) informativ belgilar majmualarining sinflashtirish samaradorligini baholash.

s) qo'yilgan masalalarni yechish algoritmlari

1-masalani yechish: Berilgan o'quv tanlamalardagi sinf obyektlarini belgi qiymatlari normallashtirish jarayoni quyidagi ketma-ketlikda amalga oshiriladi:

- a) X_p sinf obyektlarini xarakterlovchi belgi qiymatlarining maksimumi $x_p^j(max)$ elementi topiladi va bu jarayon quyidagicha hisoblanadi hamda o'qiladi:

$$x_p^j(max) = \max_i x_{pi}^j, p = \overline{1, r}; j = \overline{1, N}; i = \overline{1, m_p}; \quad (1)$$

p –sinfga tegishli obyektlar belgilarining j –belgili ustunining maksimal elementi.

- 6) X_p sinfnig obyektlarini xarakterlovchi belgi qiymati shu belgi ustunida joylashgan maksimum qiymat $x_p^j(max)$ ga bo'linadi va u quyidagicha ko'rinishda yoziladi:

$$x_{pi}^j(Normal) = \frac{x_{pi}^j}{x_p^j(max)}, p = \overline{1, r}; j = \overline{1, N}; i = \overline{1, m_p}; \quad (2)$$

p –sinfga tegishli i –obyektning j -belgisi bo'yicha normallashtirilgan holati.

2-masalani yechish: Mazkur masalani yechish uchun ishlab chiqilgan “Tibbiy ma'lumotlarga dastlabki ishlov berishning Fisher tipidagi mezon elementlariga asoslangan algoritmi”dan foydalanib, X_p sinfga tegishli obyektlarning belgi qiymatlari 0 yoki 1 qiymatga o'tkazish jarayoni quyidagicha amalga oshiriladi [15]:

1-qadam. $\bar{x}_p = (\bar{x}_p^1, \bar{x}_p^2, \dots, \bar{x}_p^N)$ vektor, X_p sinflarning o'rtacha vakil obyektlari, $p = \overline{1, r}$; Uning komponentalari quyidagi formula orqali hisoblanadi:

$$\bar{x}_p^j = \frac{1}{m_p} \sum_{i=1}^{m_p} x_{pi}^j, j = \overline{1, N}. \quad (3)$$

2-qadam. Vektorlar $a_p = (a_p^1, a_p^2, \dots, a_p^N)$ va $b_p = (b_p^1, b_p^2, \dots, b_p^N)$, komponentalari quyidagi ko'rinishda hisoblansin:

$$a_p^j = \frac{1}{m_p} \sum_{i=1}^{m_p} (\bar{x}_p^j - x_{pi}^j)^2, \quad j = \overline{1, N}. \quad (4)$$

$$b_{pi}^j = (\bar{x}_p^j - x_{pi}^j)^2, \quad j = \overline{1, N}; \quad i = \overline{1, m_p}. \quad (5)$$

3-qadam. O'quv tanlanmadagi X_p sinflardagi obyektlarning belgi qiymatlari uzluksiz miqdordan 0 yoki 1 qiymatga o'tkaziladi.

$$x_{pi}^j = \begin{cases} 1, & \text{agar } \frac{b_{pi}^j}{a_p^j} \leq 1 \\ 0, & \text{aks holda} \end{cases} \quad (6)$$

bu yerda $j = \overline{1, N}; \quad i = \overline{1, m_p}$. ga teng.

3-masalani yechish: Mazkur masala yechish ishlab chiqilgan "Etalon o'quv tanlamani shakllantirishning sinflashtirishga asoslangan algoritmi" asosida quyidagi qadamlarda amalga oshiriladi [3], [15]:

1-qadam. X_p sinfga tegishli bo'lgan har bir obyekt o'z sinfidagi hamda boshqa sinflardagi obyektlar bilan qiyoslash x_{pi} obyekt va X_p sinfnining x_{pq} obyektlari o'rtasida yaqinlik funksiyasi $\rho^j(x_{pi}, x_{pq})$ ni kiritish orqali amalga oshiriladi:

$$\rho^j(x_{pi}, x_{pq}) = \begin{cases} 1, & \text{agar } (x_{pi}^j - x_{pq}^j) = 0; \\ 0, & \text{aks holda} \end{cases} \quad (7)$$

bu yerda $j = \overline{1, N}; \quad i, q = \overline{1, m_p}; \quad i \neq q$. ga teng.

2-qadam. Qiyosiy baholash har bir sinf bo'yicha hisoblanib, ya'ni, berilgan X_p sinfnining har bir obyekt x_{pi} uchun $\Gamma_p(x_{pi}, x_{pq})$ hisoblanadi.

$$\Gamma_p(x_{pi}, x_{pq}) = \frac{1}{m_p} \sum_{q=1}^{m_p} \sum_{j=1}^N \rho^j(x_{pi}, x_{pq}), \quad i = \overline{1, m_p}; \quad i \neq q. \quad (8)$$

3-qadam. Qiyosiy baholashda olingan natijalar asosida maksimizatsiya masalasi quyidagi formula orqali yechiladi

$$\Gamma_p^*(x_{pi}, x_{pq}) = \max_{p=\overline{1, r}} \Gamma_p(x_{pi}, x_{pq}), \quad i, q = \overline{1, m_p}; \quad i \neq q. \quad (9)$$

va natijaga ko'ra $x_{pi} \in X_p$ tegishliligi aniqlanadi.

4-qadam. Noto'g'ri tanib olgan obyektlarni o'quv tanlamasidan chiqarib tashlash va 1-qadamga qaytarish.

Sinflashtirish jarayoni X_p sinf obyektlari o'z sinfini to'liq to'g'ri tanib olgunga qadar davom ettiriladi va jarayon so'ngida hosil bo'lgan tanlamani etalon o'quv tanlamasi sifatida qaraladi.

4-masalani yechish: Ushbu masalani yechish uchun ishlab chiqilgan “Sinflashtirishdagi xatolik koeffitsiyentlari va belgilarni ifodalashdagi ehtimollik ko'rsatkichlari asnosida informativ belgilar majmuasini tanlash” algoritmi asosida amalga oshiriladi [4], [15].

Unga ko'ra berilgan etalon o'quv tanlanma $x_{p1}, x_{p2}, \dots, x_{pm_p} \in X_p, p = \overline{1, r};$ uchun informativ belgilar majmuasini tanlash masalasida qism fazoni bir qiymatli xarakterlovchi $\lambda = (\lambda^1, \lambda^2, \dots, \lambda^N), \lambda^j \in \{0, 1\}, j = \overline{1, N};$ vektor kiritiladi. Bu yerda λ vektorining asosiy vazifasi N –o'lchovli belgilar fazosidan, N sonidan yetarlicha kichik bo'lgan ℓ o'lchovli belgilar fazosiga o'tishni ta'minlash hisoblanadi. Bu yerda λ vektorning birga teng bo'lgan komponentalariga mos keluvchi belgilar ajratib olinayotgan qism fazoda belgilar majmuasini bildirsa, nolga teng bo'lgan belgilari esa unga mos belgilarni ajratib olinayotgan informativ belgilar majmuasida ishtirok etmasligini bildiradi.

1-ta'rif. Tadqiq etilayotgan λ vektorlar fazosi ℓ o'lchovli deyiladi, agarda, qaralayotgan ixtiyoriy λ vektor uchun, uning komponentalari yig'indisi $\sum_{j=1}^N \lambda^j = \ell$ bo'lsa. Xuddi shuningdek, ℓ o'lchovli vektorlar fazosida, λ vektorlar joylashgan to'plam Λ^ℓ orqali belgilansin. Ta'rifga ko'ra bu to'plamning matematik ifodalanishi quyidagicha bo'ladi:

$$\Lambda^\ell = \{ \lambda: \sum_{j=1}^N \lambda^j = \ell, \lambda^j \in \{0, 1\}, j = \overline{1, N}; \}.$$

bu yerda, Λ^ℓ to'plamdagi λ vektorlar soni $C_N^\ell = \frac{N!}{\ell!(N-\ell)!}$ ga teng.

Agar $\ell = \overline{1, N}$ ta holat qaralsa, u holda bu to'plamda $2^N - 1$ ta λ vektorlar joylashgan bo'ladi, ya'ni $\sum_{\ell=1}^N C_N^\ell = 2^N - 1$.

2-ta'rif. Λ^ℓ to'plam elementlari bo'lgan λ vektorlarni ℓ informativ vektorlar deb ataladi. Bu vektorlarning noldan farqli komponentalari soni ℓ tani tashkil etadi.

Faraz qilaylik $\theta(N)$ etalon jadval asosida o'quv tanlanma obyektlari sinflashtirishdagi xatolik va noto'g'ri tanib olingan obyektlar soni $\varkappa(N)$ bo'lsin.

Bu yerda o'quv tanlanmalardagi umumiy obyektlar soni esa $M = \sum_{p=1}^r m_p$ ni tashkil etadi, u holda N o'lchovli fazoda sinflashtirishdagi xatolik quyidagi formula asosida $\theta(N) = \frac{\kappa(N)}{M}$ hisoblanadi.

Ishlab chiqilgan algoritmning ishlash mexanizmi quyidagicha aniqlangan bo'lsin. Dastlab o'quv tanlanma obyektlari belgilarini tanlashning ehtimollik vektori $p_v = (p_v^1, p_v^2, \dots, p_v^N)$ kiritiladi. Bu yerda p_v ehtimollik vektorining indeksidagi v -ehtimollik vektori ekanligini bildiradi. Odatda, bu indeks ehtimollik vektorini tanlash bilan bog'liq bo'ladi. Birinchi tanlovda $v = 1$ bo'lib, $p_1^j = \frac{1}{N}; j = \overline{1, N}$; deb olinadi. Xuddi shuningdek, $p_v = (p_v^1, p_v^2, \dots, p_v^N)$ ehtimollik vektorining ℓ -informativ $\lambda \in \Lambda^\ell$ vektorlar fazosida $p_v|_\lambda = (\lambda^1 p_v^1, \lambda^2 p_v^2, \dots, \lambda^N p_v^N)$ ko'rinishda ifodalanadi. Dastlabki berilgan ehtimollikka asosan Λ^ℓ to'plamdan $p_v = (p_v^1, p_v^2, \dots, p_v^N)$ ehtimollik vektoriga λ vektor tasodifan tanlanadi va u quyidagicha belgilanadi $\lambda_{v1}, \lambda_{v2}, \dots, \lambda_{vk}$.

Berilgan o'quv tanlanmaning barcha belgilari kesimida sinflashtirish jarayoni yuqoridagi (7), (8) va (9) formulalari asosida amalga oshiriladi. Unga ko'ra etalon o'quv tanlamasidagi barcha belgilarning alohida-alohida muhimlik darajalari aniqlanadi. Bunda etalon o'quv tanlamadan $p_v = (p_v^1, p_v^2, \dots, p_v^N)$ ehtimollik vektoridan foydalanib bitta belgi joylashgan ustun o'quv tanlanmadan chiqazib tashlab yuboriladi va qolgan belgilar kesimida sinflashtirishdagi xatolik koeffitsiyenti $\theta(N - 1)$ hisoblanadi.

Taklif etilayotgan algoritm tavsifi

1-qadam. Dastlabki sinflashtirish xatoligi $\theta(N)$ hisoblanadi;

2-qadam. ℓ ni qiymati kiritiladi, $\ell \ll N$;

3-qadam. $v = 1$ da $p_v = (p_v^1, p_v^2, \dots, p_v^N)$ uchun $p_v^j = \frac{1}{N}; j = \overline{1, N}$; qiymatlar o'zlashtiriladi;

4-qadam. Tasodifiy ravishda N ta belgilar ichidan bitta belgi tanlanadi va u majmuadan chiqazib tashlanadi. So'ngra, $N - 1$ ta belgilar kesimida sinflashtirish jarayoni amalga oshiriladi. Natijada $N - 1$ ta belgilar uchun xatolik koeffitsiyenti $\theta(N - 1)$ hisoblanadi;

5-qadam. Agar $\theta(N) = \theta(N - 1)$ bo'lsa, u holda bu belgi o'quv tanlanmadan batamom olib tashlanadi, aks holda bu belgi o'quv tanlanmada qoldiriladi va 4- qadamga o'tiladi. Bu jarayon $N - h = \ell$ bo'lguniga qadar qaytariladi.

Agar $N - h = \ell$ bo'lsa, u holda ℓ ta belgili yechim shakllantiriladi va 6- qadamga o'tiladi. Xuddi shuningdek, tenglik bajarilmasa, u holda barcha belgilar birma-bir ko'rib chiqulgunga qadar jarayon davom etadi. Barcha belgilar qo'rib chiqilgandan keyin esa, keyingi qadamga o'tiladi;

6-qadam. Faraz qilaylik, $N - h = \ell$ bo'lsin, u holda quyidagi berilgan

$$p^j = \frac{1}{2^{N-h}}, j = \overline{1, h}; p^j = \frac{2}{2^{N-h}}, j = \overline{h+1, N}; \quad \text{ehtimollik bilan 4-}$$

qadamdagi jarayon qaytariladi, faqatgina tasodifan tanlanishi lozim bo'lgan oldin tanlanib tizimda qolgan va tizimda qolmagan belgilar aro ehtimolliklar bir-biridan farq qiladi. So'ngra 5-qadamga o'tiladi.

Ushbu algoritm asosida sinflashtirish sifatini saqlagan holda sinf obyektlarini xarakterlovchi belgilarning maksimal darajada kamaytirish jarayoni amalga oshirilganda Faringit kasalligida 18 tadan 6 tagacha, Gepatit kasalligida 19 tadan 4 tagacha kamaytirishga erishildi va olingan natijalar quyidagi 3-jadvalda ifoda etilgan.

3-jadval

**Faringit va Gepatit kasalliklari uchun tanlab olingan
informativ belgilar majmualari**

Faringit ($l = 6$)		Gepatit ($l = 4$)			
1.	$x^2, x^3, x^4, x^{13}, x^{14}, x^{16}.$	1.	$x^2, x^5, x^{15}, x^{17}.$	6.	$x^1, x^2, x^5, x^{15}.$
2.	$x^3, x^4, x^9, x^{14}, x^{16}, x^{18}.$	2.	$x^2, x^5, x^{15}, x^{16}.$	7.	$x^2, x^5, x^{13}, x^{15}.$
3.	$x^3, x^4, x^{12}, x^{14}, x^{16}, x^{18}.$	3.	$x^2, x^3, x^5, x^{15}.$	8.	$x^2, x^5, x^7, x^{15}.$
4.	$x^3, x^4, x^{13}, x^{14}, x^{16}, x^{18}.$	4.	$x^2, x^5, x^{12}, x^{15}.$	9.	$x^2, x^5, x^{11}, x^{15}.$
5.	$x^3, x^4, x^{11}, x^{14}, x^{16}, x^{18}.$	5.	$x^2, x^4, x^5, x^{15}.$		

5-masalani yechish: Ushbu masalani yechish uchun Faringit va Gepatit kasalliklari tanlangan informativ belgilar majmualarini baholash uchun ishlab

chiqilgan “Informativ belgilar majmualarini baholashning ehtimollikka asoslangan imtihon qiluvchi algoritmi” algoritm asosida amalga oshiriladi [2]:

1-qadam. Etalon o‘quv tanlamani kiritish.

2-qadam. Informativ belgilar majmualarini kiritish.

3-qadam. κ soni kiritiladi. Bu qadamda etalon o‘quv tanlanma κ marta o‘quv va nazorat tanlanmaga tasodifiy ravishda ajratiladi. Bunda 90% o‘quv va 10% nazorat tanlanmalar soni κ ta bo‘ladi.

4-qadam. Informativ belgilar majmualari asosida sinflashtirish (Sinflashtirish jarayonida etalon o‘quv tanlamasidagi barcha belgilar majmuasidan faqat informativ belgilar majmualari ishtirok etadi, qolgan belgilar tashlab yuboriladi).

Mazkur qadamda etalon o‘quv tanlanma tasodifiy ravishda o‘quv va nazorat tanlanma obyektlariga ajratilgan bo‘lib, bunda tanlab olingan nazorat tanlamadagi har bir obyekt noma’lum obyekt sifatida qaralib, unga ko‘ra sinflashtirish jarayoni quyidagi matematik amallar asosida amalga oshiriladi:

a) X_p sinfga tegishli nazorat tanlamadagi har bir obyekt sinflardagi obyektlar bilan qiyoslash x_{pi} obyekt va X_p sinfnining x_{pq} obyektlari o‘rtasidagi yaqinlik funksiyasi $\rho^j(x_{pi}, x_{pq})$ ni, ya’ni (7) formulani kiritish orqali amalga oshiriladi;

b) Qiyosiy baholash har bir sinf bo‘yicha hisoblanadi. Bunda berilgan X_p sinfnining har bir obyekt x_{pi} uchun $\Gamma_p(x_{pi}, x_{pq})$, ya’ni (8) formula asosida hisoblanadi;

s) Qiyosiy baholashda olingan natijalar asosida maksimizatsiya masalasi (9) formulaga muvofiq yechiladi va natijaga ko‘ra $x_{pi} \in X_p$ tegishliligi aniqlanadi.

Natijada nazorat tanlamasidagi obyektlarning tanlangan informativ belgilar majmualari asosida sinflashtirish samaradorligi, ya’ni o‘z sinfini to‘g‘ri tanib olish sifati va xatoligini aniqlab borish orqali baholandi.

Mazkur algoritmnı tatbiq etish orqali olingan natijalar quyidagi diagrammada ifodalanilgan (4-jadval).

4-jadval

Faringit kasalligi uchun $l = 6$ dan 5 ta variantda tanlab olingan informativ belgilar majmualarining sinflashtirish samaradorligini baholash natijalari

Informativ belgilar majmualari	Tanib olish sifati (%)	Xatolik darajasi (%)
1.	98,915	1,09
2.	98,96	1,04
3.	98,845	1,16
4.	98,825	1,18
5.	98,96	1,04
O'rtacha sinflashtirish samaradorlik ko'rsatkichi(%)	98,90	1,10

5-jadval

Gepatit kasalligi uchun $l = 4$ dan 9 ta variantda tanlab olingan informativ belgilar majmualarining sinflashtirish samaradorligini baholash natijalari

Informativ belgilar majmualari	Tanib olish sifati (%)	Xatolik darajasi (%)
1.	98,95	1,06
2.	98,77	1,23
3.	99,31	0,70
4.	98,42	1,59
5.	98,87	1,13
6.	97,80	2,20
7.	98,87	1,13
8.	99,93	0,07
9.	98,91	1,10
O'rtacha sinflashtirish samaradorlik ko'rsatkichi(%)	98,87	1,13

Yuqorida keltirilgan natijalar tahlil etilganda Faringit va Gepatit kasalliklari uchun tanlangan informativ belgilar majmualarining sinflashtirish samaradorligi 98 foizdan yuqoriligini ko'rsatmoqda. Shu boisdan mazkur kasalliklar uchun tanlangan informativ belgilar majmualari, ya'ni Faringit kasalligida 5 nafar shifokor 6 ta belgidan, Gepatit kasalligida 9 nafar shifokor 4 ta belgidan foydalanib

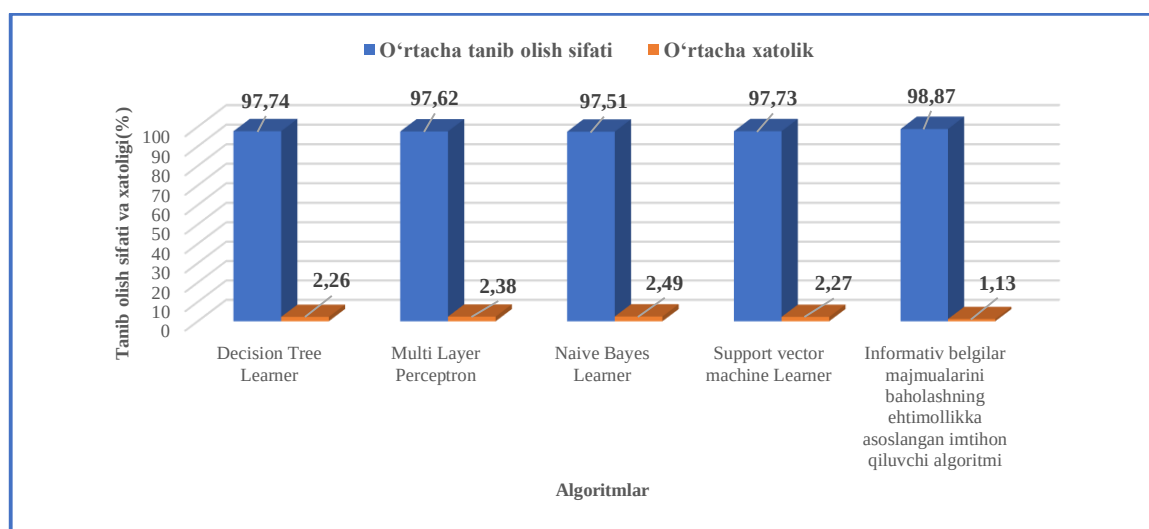
100 foizlik tashxisni qo'yayotganligini ko'rsatmoqda. Demak, shifokorlar tomonidan ishlatilgan barcha belgilar, ya'ni mazkur kasalliklar uchun bu informativ simptokomplekslar guruhini tashkil etadi va ularning ifodasi quyidagi 6-jadvalda keltirilgan.

6-jadval

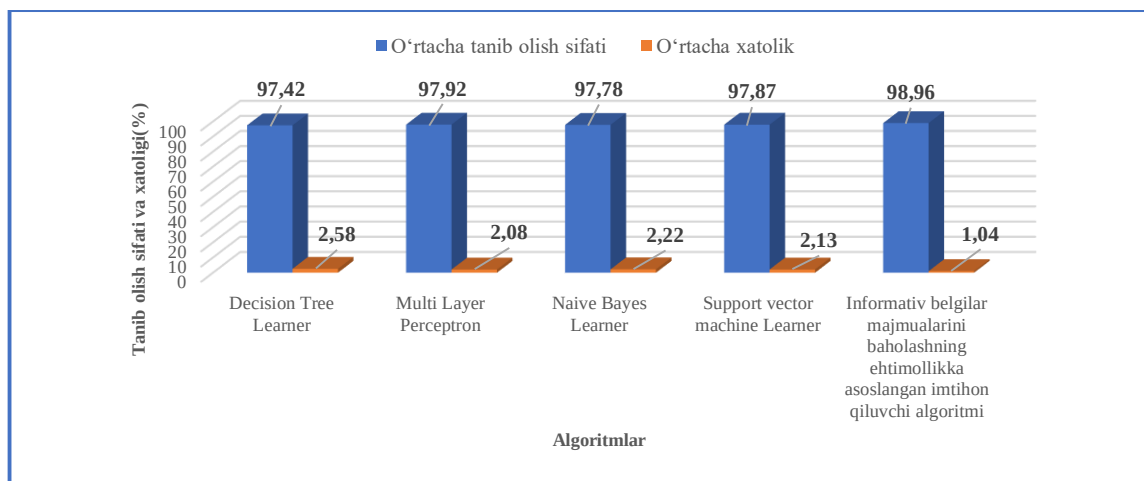
**Faringit va Gepatit kasalliklari uchun shakllantirilgan informativ
simptokomplekslar**

Faringit	Gepatit
$x^2, x^3, x^4, x^9, x^{11}, x^{12}, x^{13}, x^{14}, x^{16}, x^{18}.$	$x^1, x^2, x^3, x^4, x^5, x^7, x^{11}, x^{12}, x^{13}, x^{15}, x^{16}, x^{17}.$

Yuqorida qo'yilgan 5 ta masala va ularni yechish algoritmlari asosida ijobiy natijalar olindi. So'ngra mazkur olingan natijalarni, ya'ni Faringit kasalligi uchun $l=6$ dan 5 ta variantda, Gepatit kasalligi uchun $l = 4$ dan 9 ta variantda tanlab olingan informativ belgilar majmualarining sinflashtirish samaradorligi boshqa mashhur KNIME analitik platformasidagi mavjud algoritmlaridan foydalanilgan holda natijalar olindi va solishtirma tahlil etildi(1-2-rasmlarga qarang).



1-rasm. Faringit kasalligi uchun tanlab olingan informativ belgilar majmualarining sinflashtirish samaradorligini mavjud algoritmlardan olingan natijalar tahlili.



2-rasm. Gepatit kasalligi uchun tanlab olingan informativ belgilar majmualarining sinflashtirish samaradorligini mavjud algoritmlardan olingan natijalar tahlili.

Xulosa qilib aytganda, olingan natijalarning tahlili shuni ko'rsatmoqda:

birinchidan, ishlab chiqilgan algoritmlar asosida sinflashtirish sifatini saqlagan holda sinf obyektlarini xarakterlovchi belgilarning maksimal darajada kamaytirish jarayoni amalga oshirilganda Faringit kasalligida 18 tadan 6 tagacha, Gepatit kasalligida 19 tadan 4 tagacha kamaytirishga erishildi;

ikkinchidan, Faringit kasalligi uchun $l = 6$ dan 5 ta variantda, Gepatit kasalligi uchun $l = 4$ dan 9 ta variantda tanlab olingan informativ belgilar majmualarining sinflashtirish samaradorligi ishlab chiqilgan "Informativ belgilar majmualarini baholashning ehtimollikka asoslangan imtihon qiluvchi algoritmi" algoritmi hamda boshqa mashhur KNIME analitik platformasidagi mavjud algoritmlaridan olingan natijalar tahliliga ko'ra 97 foizdan yuqoriligi bu tanlangan informativ belgilar majmualarining salmog'i va ishonchiligi yuqoriligini ko'rsatdi.

FOYDALANILGAN ADABIYOTLAR:

1. Журавлев Ю.И. Избранные научные труды. -М: Издательство Магистр, 1998. - 420 с.
2. Жўраев Ғ.П. Тиббий тимсолларни таниб олишда танланмалар ҳолатини баҳолашнинг самарали усули // «Муҳаммад ал-Хоразмий авлодлари», илмий-амалий ва ахборот-таҳлилий журнали, Тошкент - 3(17) /2021. 13 - 20 б.

3. Нишанов А.Х., Жўраев Ғ.П., Маълумотларга дастлабки ишлов беришда синфлаштиришнинг усул ва алгоритми // Информатика ва энергетика муаммолари: Ўзбекистон журнали, №6, 2019. 19 - 26 б.
4. Нишанов А.Х., Акбаралиев Б.Б., Жўраев Ғ.П. Юрак ишемик касалликлари симптокомплексларини танлашнинг мослашувчан тасодифий кидирувга асосланган алгоритми // Муҳаммад ал-Хоразмий авлодлари: илмий-амалий ва ахборот-таҳлилий журнали. - Тошкент - 2(12) /2020. 10 - 15 б.
5. Zhang, L., Luo, M., Liu, J., Li, Z., Zheng, Q. Diverse fuzzy c-means for image clustering //Pattern Recognition Letters Volume 130, February 2020, Pages 275-283.
6. Xiong, Y., Zuo, R. Recognizing multivariate geochemical anomalies for mineral exploration by combining deep learning and one-class support vector machine//Computers and Geosciences Volume 140, July 2020, 104484
7. Li, P., Cheng, H. Track recognition algorithm based on neural network for rail transit//International Journal of Computers and Applications 2020.
8. Gai, J., Shen, J., Wang, H., Hu, Y. A Parameter-Optimized DBN Using GOA and Its Application in Fault Diagnosis of Gearbox//Shock and Vibration, Volume 2020, 2020, Номер статьи 4294095
9. Wang, D., Tian, F., Yang, S.X., (...), Jiang, D., Cai, B. Improved deep CNN with parameter initialization for data analysis of near-infrared spectroscopy sensors/Sensors (Switzerland) Volume 20, Issue 3, 1 February 2020, Номер статьи 874, 20(3), 874, 2020.
10. Si, L., Xiong, X., Wang, Z., Tan, C. A Deep Convolutional Neural Network Model for Intelligent Discrimination between Coal and Rocks in Coal Mining Face/Mathematical Problems in Engineering Volume 2020, 2020, Номер статьи 2616510
11. Fu, S., Liu, X. A new method to solve the problem of facing less learning samples in signal modulation recognition/Eurasip Journal on Wireless Communications and Networking Volume 2020, Issue 1, 1 December 2020, 8.
12. Ala'raj, M., Majdalawieh, M., Abbod, M.F. Improving binary classification

using filtering based on k-NN proximity graphs/Journal of Big Data, Volume 7, Issue 1, 1 December 2020, 15.

13.Emary E. Zawbaa H. Hassanien A. Binary ant lion approaches for feature selection// Neurocomputing. 2016 vol: 213. DOI 10.1016/j.neucom.2016.03.101. ISSN 18728286.

14.Yong Z. Dun-wei G. Wan-qiu Z. Feature selection of unreliable data using an improved multi-objective PSO algorithm// Neurocomputing. 2016 vol: 171. DOI 10.1016/j.neucom.2015.07.057. ISSN 18728286.

15.Nishanov A.Kh., Akbaraliev B.B., Juraev G.P., Khasanova M.A., Maksudova M.Kh., Umarova Z.F. The algorithm for selection of symptom complex of ischemic heart diseases based on flexible search // Journal of Cardiovascular Disease Research, Vol. 11(2), 2020. - pp. 218 - 223.